

# A LINEAR $k$ -FOLD CHEEGER INEQUALITY

FRANKLIN KENTER AND MARY RACLIFFE

**ABSTRACT.** Given an undirected graph  $G$ , the classical Cheeger constant,  $h_G$ , measures the optimal partition of the vertices into 2 parts with relatively few edges between them based upon the sizes of the parts. The well-known Cheeger's inequality states that  $2\lambda_1 \leq h_G \leq \sqrt{2\lambda_1}$  where  $\lambda_1$  is the minimum nontrivial eigenvalue of the normalized Laplacian matrix.

Recent work has generalized the concept of the Cheeger constant when partitioning the vertices of a graph into  $k > 2$  parts. While there are several approaches, recent results have shown these higher-order Cheeger constants to be tightly controlled by  $\lambda_{k-1}$ , the  $(k-1)^{\text{th}}$  nontrivial eigenvalue, to within a quadratic factor.

We present a new higher-order Cheeger inequality with several new perspectives. First, we use an alternative higher-order Cheeger constant which considers an "average case" approach. We show this measure is related to the average of the first  $k-1$  nontrivial eigenvalues of the normalized Laplacian matrix. Further, using recent techniques, our results provide linear inequalities using the  $\infty$ -norms of the corresponding eigenvectors. Consequently, unlike previous results, this result is relevant even when  $\lambda_{k-1} \rightarrow 1$ .

## 1. INTRODUCTION

Let  $G = (V, E)$  be an undirected graph, and let  $\mathcal{L} = \mathbf{D}^{-1/2}(\mathbf{I} - \mathbf{A})\mathbf{D}^{-1/2}$  be the normalized Laplacian of  $G$  with eigenvalues  $0 = \lambda_0 \leq \lambda_1 \leq \dots \leq \lambda_{n-1}$ . It is a basic fact in spectral graph theory that  $\lambda_{k-1} = 0$  if and only if  $G$  has at least  $k$  connected components. Additionally, if  $\lambda_1 \approx 0$  then the vertices of  $G$  can be partitioned into 2 parts, nearly disconnected from one another. This is formalized through the Cheeger constant and the Cheeger inequality.

The classical Cheeger constant is defined as

$$h_G = \inf_{S \subset V(G)} h(S), \text{ where } h(S) = \frac{e(S, \overline{S})}{\min\{\text{Vol}(S), \text{Vol}(\overline{S})\}},$$

where  $e(S, \overline{S})$  is the number of edges between  $S$  and its complement, and  $\text{Vol}(S)$  is the sum of vertex degrees in  $S$ . The classical Cheeger inequality relates  $h_G$  to the first eigenvalue of the normalized Laplacian matrix, as follows.

**Cheeger Inequality** (see for example [3]). Let  $\lambda_1$  be the first nontrivial eigenvalue of a connected graph  $G$ . Then

$$\frac{\lambda_1}{2} \leq h_G \leq \sqrt{2\lambda_1}.$$

Recently, some strengthenings of Cheeger's inequality have appeared (see, for example, [2, 7, 8]). Moreover, several recent results have generalized to a so-called "higher order" Cheeger constant (see, for example [9, 10, 11]) by considering a partition of  $V(G)$  into  $k > 2$  parts. While there are several different definitions of a  $k^{\text{th}}$  order Cheeger constant, one approach is to define the  $k$ -fold cheeger constant to be

$$\hat{h}_G^{(k)} = \inf_S \max_i h(S_i)$$

where the infimum ranges over all partitions of vertices  $S = \{S_1, S_2, \dots, S_k\}$ . In this case, we have:

**Higher-Order Cheeger Inequality** (Lee, Gharan and Trevisan, [9]). Let  $\lambda_{k-1}$  be the  $(k-1)^{\text{th}}$  nontrivial eigenvalue of a connected graph  $G$ . Then

$$\frac{\lambda_{k-1}}{2} \leq \hat{h}_G^{(k)} \leq O(k^2)\sqrt{2\lambda_{k-1}}.$$

This result formally demonstrates that if  $G$  can be partitioned into  $k$  parts which are nearly disconnected from one another, then  $\lambda_k \approx 0$ . Similar results for a variant of  $h_G^{(k)}$  can be found in [10].

The Cheeger constant and associated spectral information can be used to find clusters in graphs; that is, subgraphs that are highly connected. This has been a topic of wide interest in both the mathematics and computer science literature (see, for example, [4, 5, 9, 11, 12, 13, 14], among many others).

This article expands upon previous work on the Cheeger constant in two ways. First, we work with the following new notion of a  $k$ -fold Cheeger constant. For a given partition  $\mathcal{S} = \{S_1, S_2, \dots, S_k\}$  of  $V(G)$ , define the Cheeger constant of the partition,  $h_G^{(k)}(\mathcal{S})$ , to be

$$h_G^{(k)}(\mathcal{S}) = \frac{1}{k} \sum_{i \neq j} \frac{e(S_i, S_j)}{\min\{\text{Vol}(S_i), \text{Vol}(S_j)\}}.$$

We then define the  $k^{\text{th}}$  Cheeger constant of  $G$  to be  $h_G^{(k)} = \inf_{\mathcal{S}} h_G^{(k)}(\mathcal{S})$ . Specifically, while previous work focused on generalizing the Cheeger constant using a “worst case” approach, we consider the alternative “average case” approach. That is,  $\hat{h}_G^{(k)}$  requires all sets in a partition to have a small Cheeger ratio, whereas  $h_G^{(k)}$  can be small even if a small number of the sets have a large Cheeger ratio. We here reproduce a lower bound for  $h_G^{(k)}$  that agrees with the standard Cheeger inequality when  $k = 2$ . Second, we extend upon previous work of the first author [7] which gives a *linear* upper bound at the expense of using eigenvector norms. We prove:

**Theorem 1.** *Fix a constant  $k$ . Let  $G$  be an undirected graph on  $n$  vertices, with maximum degree  $\Delta$ , and suppose there exists a constant  $\beta > 0$  such that  $\frac{\Delta}{\text{Vol}(G)} = o(n^{-\beta})$ . Let  $0 = \lambda_0 \leq \lambda_1 \leq \dots \leq \lambda_{k-1}$  be the first  $k$  eigenvalues of  $\mathcal{L}$ , with corresponding harmonic eigenvectors  $x_0, x_1, \dots, x_{k-1}$ , and suppose that  $\lambda_{k-1} \leq 1$ . Let  $\alpha = \max\{\|x_i\|_{\infty} \mid i = 1, 2, \dots, k-1\}$ , and let  $\Lambda = \frac{1}{k} \sum_{i=1}^{k-1} (1 - \lambda_i)$ . Then the  $k$ -fold Cheeger constant  $h_G^{(k)}$  satisfies*

$$\frac{1}{2} - \frac{\Lambda}{2} \leq h_G^{(k)} \leq \left[ \frac{1}{2} - \frac{1}{4k} - \frac{(k-1)\Lambda}{4\text{Vol}(G)\alpha^2} \right] (1 + o(1)).$$

In addition, if we do not have  $\lambda_{k-1} \leq 1$ , we have the following related theorem.

**Theorem 2.** *Fix a constant  $k$ . Let  $G$  be an undirected graph on  $n$  vertices, with maximum degree  $\Delta$ , and suppose there exists a constant  $\beta > 0$  such that  $\frac{\Delta}{\text{Vol}(G)} = o(n^{-\beta})$ . Let  $0 = \lambda_0 \leq \lambda_1 \leq \dots \leq \lambda_{k-1}$  be the first  $k$  eigenvalues of  $\mathcal{L}$ , with corresponding harmonic eigenvectors  $x_0, x_1, \dots, x_{k-1}$ . Let  $\alpha = \sum_{i=1}^{k-1} \|x_i\|_{\infty}$ , and let  $\Lambda = \frac{1}{k} \sum_{i=1}^{k-1} (1 - \lambda_i)$ . Then the  $k$ -fold Cheeger constant  $h_G^{(k)}$  satisfies*

$$\frac{1}{2} - \frac{\Lambda}{2} \leq h_G^{(k)} \leq \left[ \frac{1}{2} - \frac{1}{4k} - \frac{(k-1)\Lambda}{4\text{Vol}(G)\alpha^2} \right] (1 + o(1)).$$

These results have several interesting features. First, the upper bounds of previous higher-order Cheeger inequalities are generally not applicable when  $\lambda_k \gg 1/k^2$  (consider the complete graph, for example). In contrast, under mild conditions, Theorem 1 or 2 applies even if  $\lambda_k \gg 1/k^2$ . Additionally, the result demonstrates that the “average case”  $k$ -fold Cheeger constant is tightly controlled by the average of the first  $k-1$  nontrivial eigenvalues whereas the previous “worst case” approaches tightly control  $\hat{h}_G^{(k)}$  with  $\lambda_k$ . Finally, and perhaps most interesting, Theorem 1 shows that the Cheeger ratio can be elegantly bounded to within a *linear* factor of the corresponding eigenvalues when the eigenvector norms are considered.

We note that although the bound in Theorem 2 appears much weaker than that of Theorem 1, it is in fact not necessarily weaker at all. Indeed, if  $\Lambda$  is negative, that is, if the average of the first  $k-1$  nontrivial eigenvalues is greater than 1, then Theorem 2 gives a stronger bound than Theorem 1, as the term involving  $\Lambda$  will be positive in this case. Indeed, as seen in Section 5, the bound given in Theorem 2 is quite good for the complete graph  $K_n$ .

We present this article as follows. In Section 2, we give relevant background and definitions. Then, we prove the lower bounds of Theorems 1 and 2 in Section 3 and the upper bounds in Section 4. Finally, we conclude with the example of applying our result to  $K_n$  in Section 5.

## 2. PRELIMINARIES

To prove the upper bound in Theorem 1, we shall use tools from both probability theory and graph theory. To begin, we define our graph-theoretic notation.

Given a graph  $G$ , define the adjacency matrix  $\mathbf{A}$  to be the square matrix, indexed by  $V(G)$ , with  $\mathbf{A}_{u,v} = \mathbf{1}_{u \sim v}$ , the indicator of whether  $\{u, v\} \in E(G)$ . Define  $\mathbf{D}$  to be the diagonal matrix indexed by  $V(G)$  with  $\mathbf{D}_{u,u} = \deg_G(u)$ . For simplicity of notation, if the graph is understood, we write  $d_u = \deg_G(u)$ . The normalized Laplacian matrix,  $\mathcal{L}$ , is given by  $\mathcal{L} = \mathbf{D}^{-1/2}(\mathbf{I} - \mathbf{A})\mathbf{D}^{-1/2}$ . By convention, if  $G$  has an isolated vertex  $u$ , set  $(\mathbf{D}^{-1/2})_{u,u} = 0$ . The eigenvalues of  $\mathcal{L}$  will be written as  $0 = \lambda_0 \leq \lambda_1 \leq \dots \leq \lambda_{n-1}$ . For a matrix or vector  $\mathbf{X}$ , we let  $\mathbf{X}^*$  denote the conjugate transpose of  $\mathbf{X}$ .

For a subset  $S \subset V(G)$ , define  $\text{Vol}(S) = \sum_{v \in S} d_v$ . Define  $\text{Vol}(G) = \text{Vol}(V(G))$ . Write  $\mathbb{1}_S$  to denote the vector indexed by  $V(G)$  with  $\mathbb{1}_S(v) = \mathbf{1}_{v \in S}$ , the indicator of whether  $v$  is an element of  $S$ . For convenience of notation, we write  $\mathbb{1} = \mathbb{1}_{V(G)}$ . Given two subsets  $S, T \subset V(G)$ , define  $e(S, T)$  to be the number of edges with one incident vertex in  $S$  and the other incident vertex in  $T$ . It is a standard exercise in graph theory to verify the following:

**Proposition 1.** *For  $S, T$  subsets of  $V(G)$ , we have*

$$e(S, T) = (\mathbf{D}^{1/2} \mathbb{1}_S)^* (\mathbf{I} - \mathcal{L}) (\mathbf{D}^{1/2} \mathbb{1}_T).$$

Given an eigenvalue-eigenvector pair  $(\lambda, \mathbf{v})$  for  $\mathcal{L}$ , define the harmonic eigenvector corresponding to  $\lambda$  to be  $\mathbf{D}^{-1/2} \mathbf{v}$ . Harmonic eigenvectors can be a useful tool for analyzing the normalized Laplacian. More information about harmonic eigenvectors and their uses can be found in [3], for example. For any graph  $G$ , there is always an orthonormal basis of eigenvectors  $\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_{n-1}$ ; we shall assume throughout that such a basis has been chosen, and the harmonic eigenvectors used have been derived from this basis, so that the harmonic eigenvector corresponding to  $\lambda_i$  will be precisely  $\mathbf{D}^{-1/2} \mathbf{v}_i$ . We note that  $\mathbf{v}_0 = (\text{Vol}(G))^{-1/2} \mathbf{D}^{1/2} \mathbb{1}$ .

In addition, we shall require the following tools from probability theory. Given a random variable  $X$ , we use  $\mathbb{E}[X]$  to denote the expected value of  $X$ . If  $A$  is an event in a probability space, we use  $\mathbf{1}_A$  to denote the 0–1 indicator random variable for  $A$ . Given a matrix  $\mathbf{M}$  whose entries are all random variables,  $\mathbb{E}[\mathbf{M}]$  denotes the matrix of entry-wise expectation. We shall use the following result from [7]:

**Proposition 2.** *Let  $\mathbf{x} \in \mathbb{C}^n$  be a random vector whose entries are pairwise independent, and let  $\mu = \mathbb{E}[\mathbf{x}]$ . If  $\mathbf{A}$  is an  $n \times n$  symmetric matrix with  $\mathbf{A}_{ii} = 0$  for all  $i$ , then*

$$\mathbb{E}[\mathbf{x}^* \mathbf{A} \mathbf{x}] = \mu^* \mathbf{A} \mu.$$

Given a random variable  $X$ , we say  $X$  has a Bernoulli distribution with parameter  $p$  if  $\mathbb{P}(X = 1) = p$ , and  $\mathbb{P}(X = 0) = 1 - p$ . We write  $X \sim \text{Ber}(p)$ .

In addition, we shall make use of Chernoff bounds. Chernoff bounds are a class of concentration inequalities that consider sums of independent random variables. Often the variables considered are Bernoulli, though that may not be the case here. There are many versions of Chernoff bounds (see, for example, [1]); we shall use the following:

**Proposition 3.** *For  $i = 1, 2, \dots, k$ , let  $X_i$  be a nonnegative random variable with  $X_i \leq \Delta$ . Let  $S = \sum X_i$ , and let  $\mu = \mathbb{E}[S]$ . Then for any  $\varepsilon > 0$ ,*

$$\mathbb{P}(|S - \mu| > \varepsilon \mu) \leq 2 \exp\left(\frac{-\varepsilon^2 \mu}{3\Delta}\right).$$

In addition, to prove the lower bound in Theorems 1 and 2, we shall make use of the following linear algebra theorem. This result can be derived as a corollary of the Courant-Fischer Theorem, and can be found, for example, as Corollary 4.3.18 in [6].

**Theorem 3.** *Let  $\mathbf{M}$  be an  $n \times n$  Hermitian matrix with eigenvalues  $\lambda_0 \leq \lambda_1 \leq \dots \leq \lambda_{n-1}$ . Fix  $k \leq n$ , and let  $\mathcal{U}_{n,k}$  denote the set of  $n \times k$  complex matrices with orthonormal columns. Then*

$$\sum_{i=0}^{k-1} \lambda_i = \min_{\mathbf{U} \in \mathcal{U}_{n,k}} \text{tr}(\mathbf{U}^* \mathbf{M} \mathbf{U}).$$

Rephrased, this theorem states that if  $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_k$  is a collection of orthonormal vectors in  $\mathbb{C}^n$  and  $\mathbf{M}$  is a Hermitian matrix, then

$$\sum_{i=1}^k \mathbf{f}_i^* \mathbf{M} \mathbf{f}_i \geq \sum_{i=0}^{k-1} \lambda_i.$$

### 3. THE LOWER BOUND

In this section, we prove the lower bound from Theorems 1 and 2. Recall that  $\Lambda = \frac{1}{k} \sum_{i=1}^{k-1} (1 - \lambda_i)$ , the average of the largest  $k$  eigenvalues of  $\mathbf{I} - \mathcal{L}$ .

**Theorem 4.** *Given  $k \geq 2$ ,*

$$h_G^{(k)} \geq \frac{1}{2} - \frac{\Lambda}{2}.$$

*Proof.* Let  $\mathcal{S} = \{S_1, S_2, \dots, S_k\}$  be a partition of the vertices of  $G$ . For  $1 \leq j \leq k$ , let  $g_j = \mathbf{D}^{1/2} \frac{1}{\sqrt{\text{Vol}(S_i)}} \mathbb{1}_{S_i}$ . Note that as the  $S_i$  are all disjoint, we have that the set  $\{g_j\}$  is orthogonal. Moreover,  $\|g_j\|^2 = \sum_{v \in S_i} \frac{d_v}{\text{Vol}(S_i)} = 1$ , and hence the set  $\{g_j\}$  is in fact orthonormal. Then by Theorem 3, we have

$$\sum_{i=1}^k (g_i^* \mathcal{L} g_i) \geq \sum_{i=0}^{k-1} \lambda_i = k(1 - \Lambda).$$

On the other hand, by Proposition 1, we have

$$g_i^* \mathcal{L} g_i = g_i^* g_i - g_i^* (\mathbf{I} - \mathcal{L}) g_i = 1 - \frac{e(S_i, S_i)}{\text{Vol}(S_i)}.$$

Combining these two results yields

$$\begin{aligned} 1 - \Lambda &\leq \frac{1}{k} \left( \sum_{i=1}^k \left( 1 - \frac{e(S_i, S_i)}{\text{Vol}(S_i)} \right) \right) \\ &= \frac{1}{k} \left( \sum_{i=1}^k \left( 1 - \frac{e(S_i, S_i) + e(S_i, \overline{S_i})}{\text{Vol}(S_i)} + \sum_{j \neq i} \frac{e(S_i, S_j)}{\text{Vol}(S_i)} \right) \right) \\ &= \frac{1}{k} \left( \sum_{i=1}^k \left( 1 - \frac{\text{Vol}(S_i)}{\text{Vol}(S_i)} + \sum_{j \neq i} \frac{e(S_i, S_j)}{\text{Vol}(S_i)} \right) \right) \\ &= \frac{1}{k} \sum_{j \neq i} \left( \frac{e(S_i, S_j)}{\text{Vol}(S_i)} + \frac{e(S_i, S_j)}{\text{Vol}(S_j)} \right) \\ &\leq 2h_G^{(k)}, \end{aligned}$$

as desired. □

### 4. THE UPPER BOUND

The proof of the upper bound in Theorem 1 is modeled after the proof of a similar upper bound by the first author in [7] when  $k = 2$ . The strategy employed is to choose a vector randomly so that the expectation of this vector has useful algebraic properties, and apply concentration results for the expectation.

*Proof of Theorem 1, upper bound.* Let  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{k-1}$  be the first  $k - 1$  nontrivial harmonic eigenvectors for  $G$ , with corresponding eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_{k-1}$ . As above, we shall assume that  $\mathbf{x}_i = \mathbf{D}^{-1/2} \mathbf{v}_i$ , where the set  $\{\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_n\}$  is an orthonormal basis for  $\mathbb{R}^n$  composed of eigenvectors of  $\mathcal{L}$ . Recall that  $\mathbf{v}_0 = \frac{1}{\sqrt{\text{Vol}(G)}} \mathbf{D}^{1/2} \mathbb{1}$ , and hence for each  $i$ ,

$$\sum_{v \in V(G)} \mathbf{x}_i(v) d_v = \mathbf{x}_i^* \mathbf{D} \mathbb{1} = \sqrt{\text{Vol}(G)} (\mathbf{D}^{-1/2} \mathbf{v}_i)^* \mathbf{D}^{1/2} \mathbf{v}_0 = \sqrt{\text{Vol}(G)} \mathbf{v}_i^* \mathbf{v}_0 = 0,$$

so each  $\mathbf{x}_i$  is orthogonal to  $D\mathbb{1}$ . Let  $\alpha = \max\{\|\mathbf{x}_j\|_\infty \mid j = 1, 2, \dots, k\}$ .

Let  $\delta > 0$  be a constant that will be defined later. For all  $v \in V(G)$ , define a random variable  $s_v$  by

$$s_v = j \text{ with probability } \frac{1-2\delta}{2(k-1)} + \frac{\mathbf{x}_j(v)}{2(k-1)\|\mathbf{x}_j\|_\infty} \text{ for } j = 1, 2, \dots, k-1$$

$$s_v = k \text{ otherwise}$$

Note that the random variables for different vertices are independent. For each  $j = 1, 2, \dots, k$ , define  $S_j = \{v \in V \mid s_v = j\}$ . Thus we can view the random variables  $s_v$  as partitioning the vertices of  $G$  into  $k$  sets. Let  $\mathbf{w}_j = \mathbb{1}_{S_j}$ , the indicator vector for the set  $S_j$ . As the choice of set  $S_j$  for each vertex  $v$  is independent of each other vertex, we have that the entries in  $\mathbf{w}_j$  are pairwise independent (although any pair  $\mathbf{w}_j, \mathbf{w}_\ell$  are not independent). Notice that for  $j = 1, 2, \dots, k-1$ , we have

$$\begin{aligned} \mathbb{E}[\text{Vol}(S_j)] &= \sum_{v \in V} \mathbb{P}(v \in S_j) d_v \\ &= \sum_{v \in V} \left( \left( \frac{1-2\delta}{2(k-1)} \right) d_v + \frac{1}{2(k-1)\|\mathbf{x}_j\|_\infty} \mathbf{x}_j(v) d_v \right) \\ &= \frac{1-2\delta}{2(k-1)} \text{Vol}(G) =: \mu, \end{aligned}$$

by orthogonality of  $\mathbf{x}_j$  to  $\mathbf{D}\mathbf{1}$ . Note that  $\mu$  is independent of the choice of  $j$ .

On the other hand, for a given  $j \leq k-1$ , we can view the vertices in  $S_j$  as chosen by a sequence of independent Bernoulli random variables,  $X_v$ , where  $X_v \sim \text{Ber}\left(\frac{1-2\delta}{2(k-1)} + \frac{\mathbf{x}_j(v)}{2(k-1)\|\mathbf{x}_j\|_\infty}\right)$ . Thus,  $\text{Vol}(S_j) = \sum_{v \in V(G)} d_v X_v$ , and by Proposition 3, we have that for all  $\varepsilon > 0$ ,

$$\mathbb{P}(|\text{Vol}(S_j) - \mu| > \varepsilon \mu) \leq 2 \exp\left(\frac{-\varepsilon^2 \mu}{3\Delta}\right),$$

where  $\Delta$  is the maximum degree of a vertex in  $G$ .

Let  $A$  be the event that  $(1-\varepsilon)\mu < \text{Vol}(S_j) < (1+\varepsilon)\mu$  for all  $1 \leq j \leq k-1$ . By the union bound, we have

$$\mathbb{P}(A^c) \leq 2(k-1) \exp\left(\frac{-\varepsilon^2 \mu}{3\Delta}\right).$$

Thus for each  $j$ , we have

$$\begin{aligned} \mathbb{E}[e(S_j, S_j)] &= \mathbb{E}[e(S_j, S_j)\mathbf{1}_A + e(S_j, S_j)\mathbf{1}_{A^c}] \\ &\leq \mathbb{E}[e(S_j, S_j)\mathbf{1}_A] + \text{Vol}(G) \left( 2(k-1) \exp\left(\frac{-\varepsilon^2 \mu}{3\Delta}\right) \right) \\ &= \mathbb{E}\left[\frac{e(S_j, S_j)}{\text{Vol}(S_j)} \text{Vol}(S_j) \mathbf{1}_A\right] + 2(k-1) \text{Vol}(G) \exp\left(-\frac{\varepsilon^2 \mu}{3\Delta}\right) \\ &= \mathbb{E}\left[\frac{e(S_j, S_j) + e(S_j, \overline{S_j}) - e(S_j, \overline{S_j})}{\text{Vol}(S_j)} \text{Vol}(S_j) \mathbf{1}_A\right] + 2(k-1) \text{Vol}(G) \exp\left(-\frac{\varepsilon^2 \mu}{3\Delta}\right) \\ &= \mathbb{E}\left[\frac{\text{Vol}(S_j) - e(S_j, \overline{S_j})}{\text{Vol}(S_j)} \text{Vol}(S_j) \mathbf{1}_A\right] + 2(k-1) \text{Vol}(G) \exp\left(-\frac{\varepsilon^2 \mu}{3\Delta}\right). \end{aligned}$$

Using linearity of expectation, we can sum over  $j$  to obtain

$$\begin{aligned}
\mathbb{E} \left[ \sum_{j=1}^{k-1} e(S_j, S_j) \right] &\leq \sum_{j=1}^{k-1} \left[ \mathbb{E} \left[ \frac{\text{Vol}(S_j) - e(S_j, \bar{S}_j)}{\text{Vol}(S_j)} \text{Vol}(S_j) \mathbf{1}_A \right] + 2(k-1) \text{Vol}(G) \exp(-\varepsilon^2 \mu / (3\Delta)) \right] \\
&\leq (1+\varepsilon) \mu \mathbb{E} \left[ \sum_{j=1}^{k-1} \left( 1 - \frac{e(S_j, \bar{S}_j)}{\text{Vol}(S_j)} \right) \mathbf{1}_A \right] + 2(k-1)^2 \text{Vol}(G) \exp(-\varepsilon^2 \mu / (3\Delta)) \\
&= (1+\varepsilon) \mu \mathbb{E} \left[ \left( k-1 - \sum_{j=1}^{k-1} \sum_{\substack{i=1 \\ j \neq i}}^k \frac{e(S_j, S_i)}{\text{Vol}(S_j)} \right) \mathbf{1}_A \right] + 2(k-1)^2 \text{Vol}(G) \exp(-\varepsilon^2 \mu / (3\Delta)) \\
&= (1+\varepsilon) \mu \mathbb{E} \left[ \left( k-1 - \sum_{i \neq j} \left( \frac{e(S_j, S_i)}{\text{Vol}(S_j)} + \frac{e(S_i, S_j)}{\text{Vol}(S_i)} \right) + \sum_{j=1}^{k-1} \frac{e(S_j, S_k)}{\text{Vol}(S_k)} \right) \mathbf{1}_A \right] \\
&\quad + 2(k-1)^2 \text{Vol}(G) \exp(-\varepsilon^2 \mu / (3\Delta)) \\
&\leq (1+\varepsilon) \mu \mathbb{E} \left[ \left( k-1 - \sum_{i \neq j} \frac{2e(S_j, S_i)}{\min\{\text{Vol}(S_i), \text{Vol}(S_j)\}} \frac{1-\varepsilon}{1+\varepsilon} + 1 \right) \mathbf{1}_A \right] \\
&\quad + 2(k-1)^2 \text{Vol}(G) \exp(-\varepsilon^2 \mu / (3\Delta)) \\
&\leq (1+\varepsilon) \mu \left( k - 2kh_G^{(k)} \frac{1-\varepsilon}{1+\varepsilon} \right) + 2(k-1)^2 \text{Vol}(G) \exp(-\varepsilon^2 \mu / (3\Delta)).
\end{aligned}$$

On the other hand, we also have  $\mathbf{w}_j = \mathbb{1}_{S_j}$ ,

$$\mathbb{E}[\mathbf{w}_j] = \left( \frac{1-2\delta}{2(k-1)} \right) \mathbb{1} + \frac{1}{2\|\mathbf{x}_j\|_\infty} \mathbf{x}_j,$$

and thus, as the entries in  $\mathbf{w}_j$  are independent as noted above, by Proposition 2, we have

$$\begin{aligned}
\mathbb{E}[e(S_j, S_j)] &= \mathbb{E} \left[ (\mathbf{D}^{1/2} \mathbf{w}_j)^* (\mathbf{I} - \mathcal{L}) (\mathbf{D}^{1/2} \mathbf{w}_j) \right] \\
&= \left( \left( \frac{1-2\delta}{2(k-1)} \right) \mathbb{1} + \frac{1}{2\|\mathbf{x}_j\|_\infty} \mathbf{x}_j \right)^* \mathbf{D}^{1/2} (\mathbf{I} - \mathcal{L}) \mathbf{D}^{1/2} \left( \left( \frac{1-2\delta}{2(k-1)} \right) \mathbb{1} + \frac{1}{2\|\mathbf{x}_j\|_\infty} \mathbf{x}_j \right) \\
&= \left( \frac{1-2\delta}{2(k-1)} \right)^2 \text{Vol}(G) + \frac{1-\lambda_j}{4\|\mathbf{x}_j\|_\infty^2}.
\end{aligned}$$

We therefore obtain

$$\sum_{j=1}^{k-1} \left( \left( \frac{1-2\delta}{2(k-1)} \right)^2 \text{Vol}(G) + \frac{1-\lambda_j}{4\|\mathbf{x}_j\|_\infty^2} \right) \leq (1+\varepsilon) \mu \left( k - 2kh_G^{(k)} \frac{1-\varepsilon}{1+\varepsilon} \right) + 2(k-1)^2 \text{Vol}(G) \exp(-\varepsilon^2 \mu / (3\Delta)).$$

Recall that  $\frac{\Delta}{\text{Vol}(G)} = o(n^{-\beta})$  by hypothesis; choose  $\delta = \varepsilon = n^{-\beta/3}$ . Then we have

$$\exp(-\varepsilon^2 \mu / (3\Delta)) = \exp \left( -\frac{n^{-2\beta/3} (1 - 2n^{-\beta/3}) \text{Vol}(G)}{6(k-1)\Delta} \right) \leq \exp(-C(1 - o(1))n^{\beta/3}),$$

for  $C$  an appropriate constant. As  $\text{Vol}(G) \leq n^2$ , we thus have that the error term satisfies

$$2(k-1)^2 \text{Vol}(G) \exp(-\varepsilon^2 \mu / (3\Delta)) \leq 2(k-1)^2 n^2 \exp(-C(1 - o(1))n^{\beta/3}) = o(1).$$

Moreover, as  $\|\mathbf{x}_j\|_\infty \geq \alpha$  and  $1 - \lambda_j \geq 0$  for all  $j$ , we have

$$(1) \quad \sum_{j=1}^{k-1} \frac{1-\lambda_j}{4\|\mathbf{x}_j\|_\infty^2} \geq \frac{k\Lambda}{4\alpha^2}.$$

Therefore,

$$\begin{aligned} \sum_{j=1}^{k-1} \left( \left( \frac{1-2\delta}{2(k-1)} \right)^2 \text{Vol}(G) + \frac{1-\lambda_j}{4\|\mathbf{x}_j\|_\infty^2} \right) &\leq (1+\varepsilon)\mu \left( k - 2kh_G^{(k)} \frac{1-\varepsilon}{1+\varepsilon} \right) + o(1) \\ \frac{(1-2\delta)^2}{4(k-1)} \text{Vol}(G) + \frac{k\Lambda}{4\alpha^2} &\leq (1+o(1)) \frac{\text{Vol}(G)}{2(k-1)} \left( k - 2kh_G^{(k)}(1-o(1)) \right) + o(1) \end{aligned}$$

Solving for  $h_G^{(k)}$  yields

$$h_G^{(k)} \leq \left[ \frac{1}{2} - \frac{1}{4k} - \frac{(k-1)\Lambda}{4\text{Vol}(G)\alpha^2} \right] (1+o(1)),$$

as desired.  $\square$

To obtain the proof of Theorem 2, we note that the only inequality that fails above when  $\lambda_{k-1} > 1$  is (1). To correct for this problem, we shall slightly modify the definition of the random variable  $s_v$  for each  $v$ .

*Proof of Theorem 2, upper bound.* As in the proof of Theorem 1, Let  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{k-1}$  be the first  $k-1$  nontrivial harmonic eigenvectors for  $G$ , with corresponding eigenvalues  $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_{k-1}$ . Let  $\alpha = \sum_{i=1}^{k-1} \|\mathbf{x}_i\|_\infty$ . Let  $\delta > 0$ , and for all  $v \in V(G)$ , define a random variable  $s_v$  by

$$\begin{aligned} s_v &= j \text{ with probability } \frac{1-2\delta}{2(k-1)} + \frac{\mathbf{x}_j(v)}{2(k-1)\alpha} \text{ for } j = 1, 2, \dots, k-1 \\ s_v &= k \text{ otherwise} \end{aligned}$$

Proceed with the proof as in Theorem 1, noting that we can replace inequality (1) with the equality

$$\sum_{j=1}^{k-1} \frac{1-\lambda_j}{4\alpha^2} = \frac{k\Lambda}{4\alpha^2},$$

and that all else is unchanged. The result then follows.  $\square$

## 5. EXAMPLE

As an example of an application of Theorem 2, we consider the complete graph  $K_n$ . Take  $k$  to be a fixed constant, and we shall consider the asymptotics of  $h_G^{(k)}$  as  $n \rightarrow \infty$ .

As with the standard Cheeger constant, it is quite clear to see that  $h_G^{(k)}(\mathcal{S})$  will be minimized when the  $S_i$  are roughly an equipartition of  $n$ , that is, when there are exactly  $r = n \bmod k$  sets of size  $\lceil \frac{n}{k} \rceil$  and the rest are of size  $\lfloor \frac{n}{k} \rfloor$ . Letting  $\mathcal{S}$  be such a partition, we have

$$\begin{aligned} h_{K_n}^{(k)} &= \frac{1}{k} \sum_{i \neq j} \frac{|S_i||S_j|}{\min\{(n-1)|S_i|, (n-1)|S_j|\}} \\ &= \frac{1}{k(n-1)} \left( \binom{r}{2} \left\lceil \frac{n}{k} \right\rceil + \binom{k-r}{2} \left\lfloor \frac{n}{k} \right\rfloor + \left( \binom{k}{2} - \binom{r}{2} - \binom{k-r}{2} \right) \left\lceil \frac{n}{k} \right\rceil \right) \\ &\sim \frac{1}{k(n-1)} \left( \binom{k}{2} \frac{n}{k} \right) \\ &\sim \frac{1}{2} - \frac{1}{2k}. \end{aligned}$$

Recall that the Laplacian eigenvalues of  $K_n$  are  $\lambda_0 = 0$  and  $\lambda_i = \frac{n}{n-1}$  for all  $i > 0$ . Therefore, the lower bound given in Theorem 1 yields  $h_{K_n}^{(k)} \geq \frac{1}{2k}(k-1)\frac{n}{n-1} \sim \frac{1}{2} - \frac{1}{2k}$ , a true estimate for the Cheeger ratio.

Moreover,  $\Lambda = \frac{k-1}{k}(-\frac{1}{n-1})$ . Also, we have  $\mathbf{v}_0$ , the eigenvector corresponding to 0, is given by  $\mathbf{v}_0 = \frac{1}{\sqrt{n}}\mathbb{1}$ . Thus any vector perpendicular to  $\mathbf{v}_0$  is an eigenvector for  $\frac{n}{n-1}$ . Note that as  $\Lambda$  is negative here, we wish to maximize  $\alpha^2$  in order to minimize the upper bound. Thus we take  $\mathbf{v}_i = \frac{1}{\sqrt{2}}(\mathbf{e}_{2i-1} - \mathbf{e}_{2i})$ . It is clear

that these are orthonormal. In addition, this implies that the harmonic eigenvectors are  $\mathbf{x}_i = \mathbf{D}^{-1/2} \mathbf{v}_i = \frac{1}{\sqrt{2(n-1)}}(\mathbf{e}_{2i-1} - \mathbf{e}_{2i})$ , and hence  $\alpha^2 = \left(\frac{k-1}{\sqrt{2(n-1)}}\right)^2 = \frac{(k-1)^2}{2(n-1)}$ .

Therefore, we obtain as the upper bound from Theorem 2

$$\begin{aligned} h_{K_n}^{(k)} &\leq \left[ \frac{1}{2} - \frac{1}{4k} - \frac{(k-1)\Lambda}{4 \text{Vol}(G) \alpha^2} \right] (1 + o(1)) \\ &= \left[ \frac{1}{2} - \frac{1}{4k} - \frac{(k-1) \frac{k-1}{k} \left(-\frac{1}{n-1}\right)}{4n(n-1) \frac{(k-1)^2}{2(n-1)}} \right] (1 + o(1)) \\ &= \left( \frac{1}{2} - \frac{1}{4k} \right) (1 + o(1)), \end{aligned}$$

compared to the true constant  $\frac{1}{2} - \frac{1}{2k}$ .

## REFERENCES

- [1] N. ALON AND J. SPENCER, *The Probabilistic Method*, John Wiley & Sons, 3 ed., 2008.
- [2] F. R. CHUNG, *Laplacians of graphs and cheeger's inequalities*, in Proc. Int. Conf. Combinatorics, Paul Erdos is Eighty, Keszthely (Hungary), vol. 2, 1993, p. 116.
- [3] ———, *Spectral graph theory*, vol. 92, American Mathematical Soc., 1997.
- [4] M. GIRVAN AND M. E. NEWMAN, *Community structure in social and biological networks*, Proceedings of the National Academy of Sciences, 99 (2002), pp. 7821–7826.
- [5] F. C. GRAHAM AND A. TSIATAS, *Finding and visualizing graph clusters using pagerank optimization*, in Algorithms and Models for the Web-Graph, Springer, 2010, pp. 86–97.
- [6] R. HORN AND C. JOHNSON, *Matrix Analysis*, Cambridge University Press, 1985.
- [7] F. KENTER, *A linear cheeger inequality using eigenvector norms*, Journal of Combinatorics, (to appear).
- [8] T. C. KWOK, L. C. LAU, Y. T. LEE, S. OVEIS GHARAN, AND L. TREVISAN, *Improved cheeger's inequality: analysis of spectral partitioning algorithms through higher order spectral gap*, in Proceedings of the forty-fifth annual ACM symposium on Theory of computing, ACM, 2013, pp. 11–20.
- [9] J. R. LEE, S. OVEIS GHARAN, AND L. TREVISAN, *Multi-way spectral partitioning and higher-order cheeger inequalities*, in Proceedings of the forty-fourth annual ACM symposium on Theory of computing, ACM, 2012, pp. 1117–1130.
- [10] A. LOUIS, P. RAGHAVENDRA, P. TETALI, AND S. VEMPALA, *Algorithmic extensions of cheeger's inequality to higher eigenvalues and partitions*, in Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, Springer, 2011, pp. 315–326.
- [11] ———, *Many sparse cuts via higher eigenvalues*, in Proceedings of the forty-fourth annual ACM symposium on Theory of computing, ACM, 2012, pp. 1131–1140.
- [12] A. Y. NG, M. I. JORDAN, Y. WEISS, ET AL., *On spectral clustering: Analysis and an algorithm*, Advances in neural information processing systems, 2 (2002), pp. 849–856.
- [13] S. E. SCHAEFFER, *Graph clustering*, Computer Science Review, 1 (2007), pp. 27–64.
- [14] D. A. SPIELMAN AND S.-H. TENG, *A local clustering algorithm for massive graphs and its application to nearly-linear time graph partitioning*, arXiv preprint arXiv:0809.3232, (2008).